Kabayaki for Linux

取扱説明書

version 1.2.0

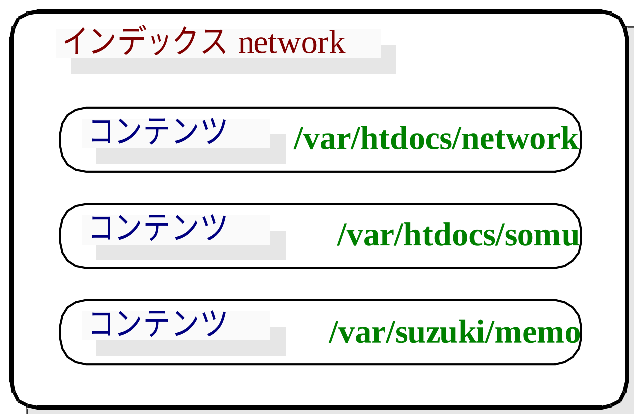
目次

はじめに

Kabayaki の管理とは

コンテンツとインデックス

Kabayaki は，コンテンツとインデックスいう仕組みを使って文書の検索をおこないます．コンテンツとインデックスとは次のような関係になります．



ここでは，Kabayaki を運用するシステム上のファイルシステムで，/var/htdocs/network ，/var/htdocs/somu ，/home/suzuki/memo という，別々の場所で管理している社内ネットワーク関連文書を， network という１つのインデックスにまとめて管理する場合を例にとります．

インデックスとコンテンツをうまく使うことにより，複数のディレクトリに散在しているファイルをまとめて管理することができます．

１章
管理画面の構成

管理画面の表示

Kabayaki の管理は，Microsoft Internet Explorer や Netscape Communications Netscape Navigator などのウェブブラウザを使い，管理画面を表示させておこないます．管理画面を表示させるには，ウェブブラウザへ次のように URL を入力します．

http://<ホスト名>/kabayaki/admin/rc.cgi

<ホスト名> には，Kabayaki をインストールしたコンピュータの名前を充てます．インストールに関しては，別冊の『インストール取扱説明書』をご覧ください．
たとえば，search.timedia.co.jp というホストに Kabayaki をインストールした場合は，次のようになります．

http://search.timedia.co.jp/kabayaki/admin/rc.cgi

続いて，ウェブブラウザに次のような Kabayaki 管理画面が表示されます．

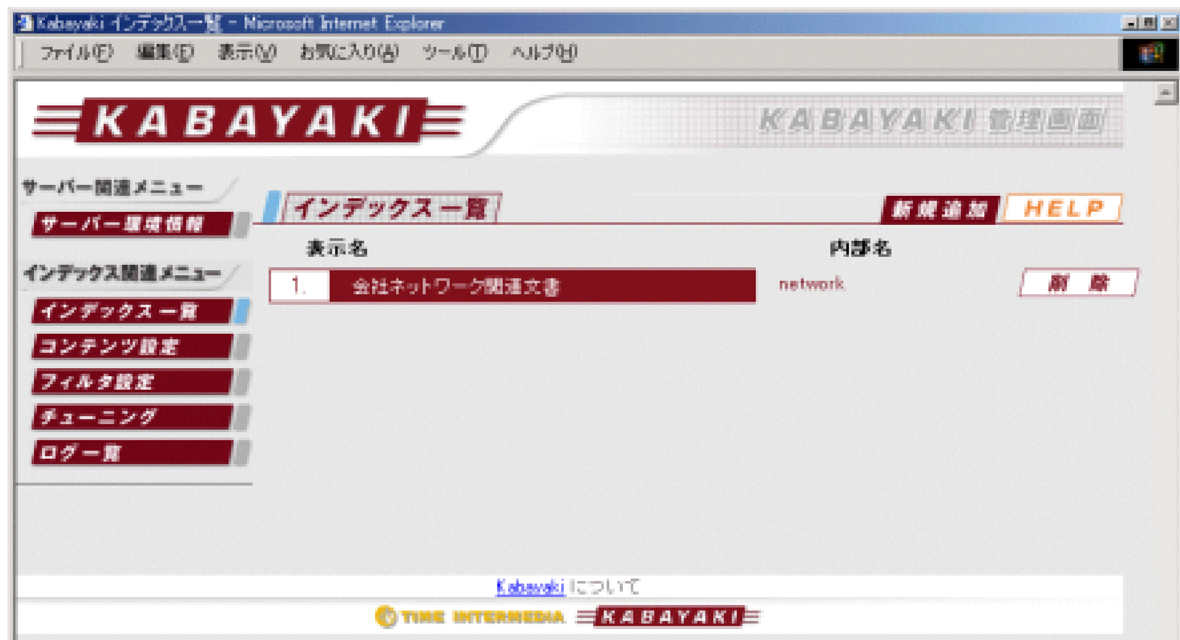


管理画面左側の一覧をメインメニューと呼びます．メインメニューは，サーバ関連メニュー，インデックス関連メニューから構成され，オプションパックをインストールしている場合は，さらに一覧が追加されます．オプションパックに関しては，オプションパックに付属の取扱説明書をご覧ください．

マウスカーソルを各メニューの一覧へ重ねると，インジケータと呼ばれる部分の色が変化します．メニューから選択するには，インジケータが変化した一覧をクリックします．

サーバ関連メニューでは，Kabayaki をインストールしたホスト全体に関わる情報の設定や表示をおこなう機能を選べます．インデックス関連メニューでは，検索に必要なインデックスの情報の設定や表示をおこなう機能を選べます．メインメニューは，機能を選んで画面が切り替わっても，つねに表示されています．

Kabayaki をインストールして，最初に Kabayaki 管理画面を表示させたときは，サーバー関連メニューのサーバー環境情報とインデックス関連メニューのインデックス一覧しか選ぶことができません．その他の一覧や設定の選び方はあとで解説します，インデックスを作成しなければなりません．



さて，各設定画面に共通で表示されるものに，さらに次のものがあります．

[新規追加］ボタン

ページ右上に表示されます．クリックすると，新たにインデックスを追加するための設定画面が表示されます．インデックス関連メニューを選択したときのみ表示されます．

[HELP］ボタン

ページ右上に表示されます．クリックすると，各設定画面ごとのオンラインヘルプが表示されます．

[削除］ボタン
ページ右側やや上部に，選択されているインデックスがあるときのみ表示されます．クリックすると，選択されているインデックスが削除されます.
[インデックス選択］リスト
ページ中央やや上部に，クリックすると一覧が表示されます．一覧から一つをクリックすると，インデックスが変更されます．インデックス関連メニューを選択したときのみ表示されます．インデックスがある場合に情報がページ上部に表示され，インデックスの選択や削除ができます.
[保存］ボタン
ページ下部に，クリックすると，ページに入力されている内容をホストへ保存します.
さて，次に，各関連メニューについて解説します.

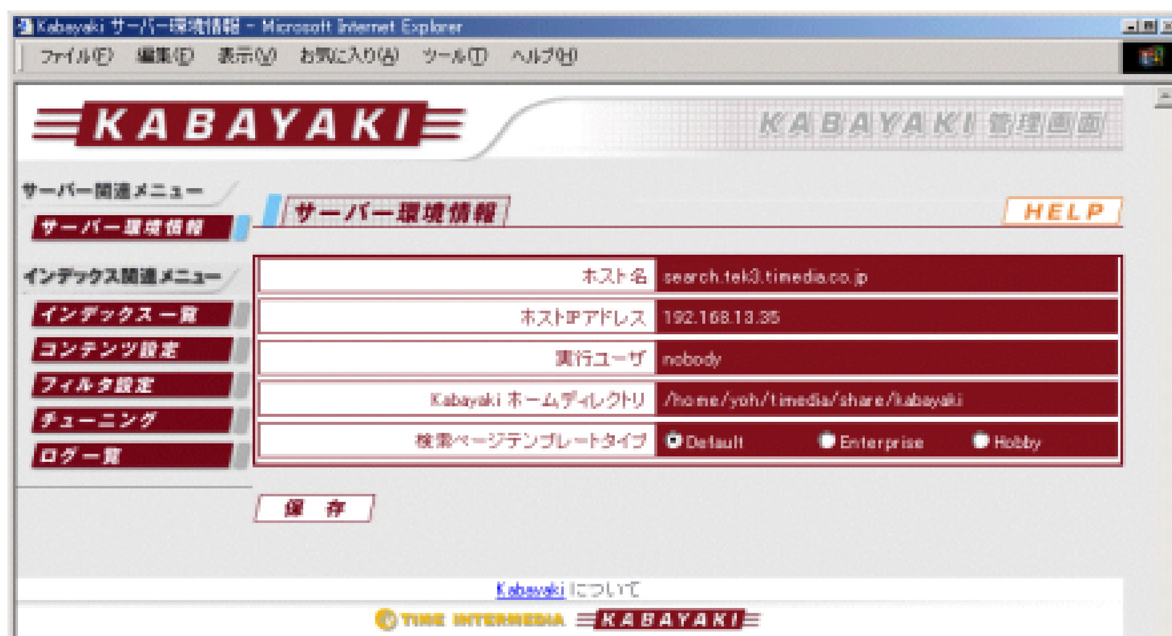
2章
サーバー関連
サーバー環境

サーバー関連メニュー

サーバー環境情報

Kabayaki をインストールしたホストに関わる情報を表示し，検索と検索結果の画面のデザインを変更します．

サーバー環境情報メニューをクリックすると，次のような画面が表示されます．



画面に表示されている情報は以下のとおりです．

ホスト名

Kabayaki が動作しているホストの名前が表示されます．これは，環境変数 SERVER_NAME を参照しています．環境変数については，ホストのシステム管理者にお尋ねください．なお，SERVER_NAME が設定されていない場合は，`不明`と表示されます．

ホスト IP アドレス

Kabayaki が動作しているホストの IP アドレスが表示されます．これは，環境変数 SERVER_ADDR を参照しています．なお，SERVER_ADDR が設定されていない場合は，`不明`と表示されます．

実行ユーザ

Kabayaki 管理画面を実行しているプロセスの実行ユーザ名が表示されます．プロセスと実行ユーザ名については，ホストのシステム管理者にお尋ねください．

Kabayaki ホームディレクトリ

インストールされている Kabayaki の先頭のディレクトリを表示します．これは，環境変数 KABAYAKI_HOME を参照しています．

検索ページテンプレートタイプ

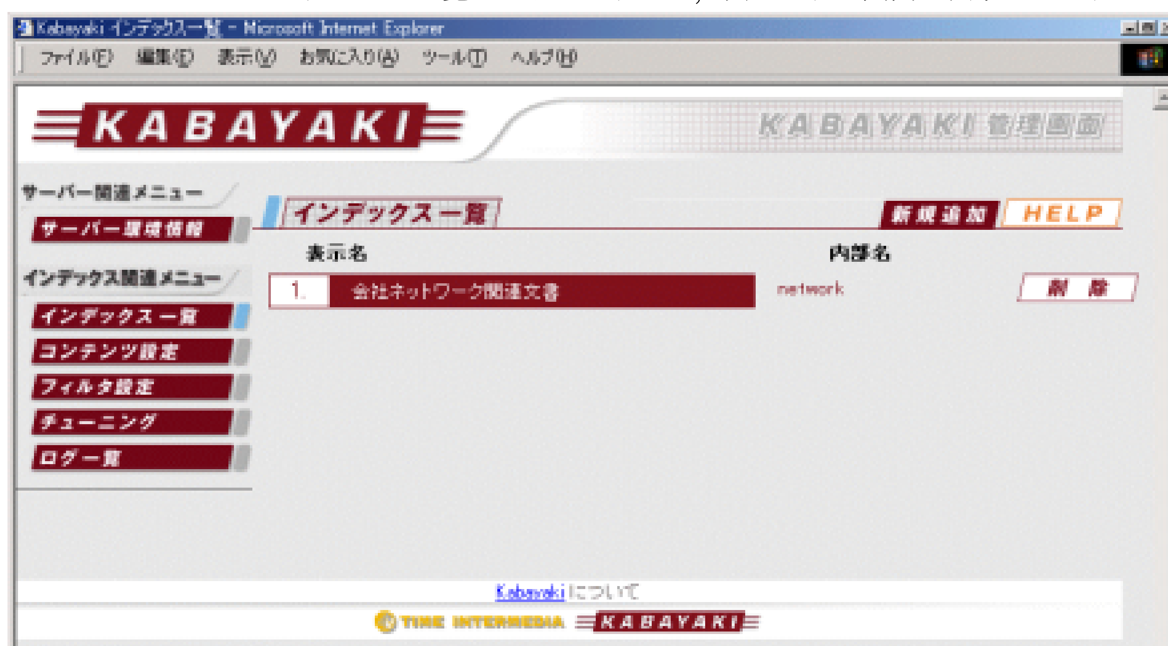
検索と検索結果の画面のようすを選択できます．Default は，インストールした直後と同じ画面，Enterprise は，高級感のある雰囲気の画面，Hobby は遊び心のある画面になります．

［保存］ボタン

［保存］ボタンをクリックすると，検索ページテンプレートタイプの設定を保存します.

3章
インデックス関連
一覧・フィルタなど

インデックス関連メニュー
インデックス一覧
メインメニューのインデックス一覧をクリックすると，次のような画面が表示されます．



登録されているインデックスの一覧を表示します．Kabayaki をインストールした直後と登録されているインデックスが一つも無い場合は，

新しくインデックスを作成してください。

と表示されます．なお，インデックスについては，『Kabayaki の管理とは コンテンツとインデックス』を参照してください．

インデックス一覧の表示は，あとで解説します内部名のアルファベット順に並び替えられます．一覧で表示される項目は次の通りです．

表示名
表示名には，あとで解説します新規追加の際，ユーザが判読しやすい，インデックスに対応した名前を付けることができます．クリックすると，インデックスのコンテンツ設定ページが表示されます．

内部名
内部名は，Kabayaki が利用する名前で，新規追加の際，ユーザが使うことのできる文字は限られています．クリックすると，インデックスのコンテンツ設定ページが表示されます．

さて，新しくインデックスを追加するには，［新規追加］ボタンをクリックします．次のような新規インデックス追加ページが表示されます．

内部名

インデックスを識別するために Kabayaki が利用する名前を入力します．半角小文字の英数字とアンダースコア

0123456789 abcdefghijklmnopqrstuvwxyz _

のみが入力できます．

表示名

Kabayaki 管理画面や検索結果画面で表示されるインデックス名を入力します．内部名と違い，入力できる文字に制限はありませんので，判りやすいものを選べます．

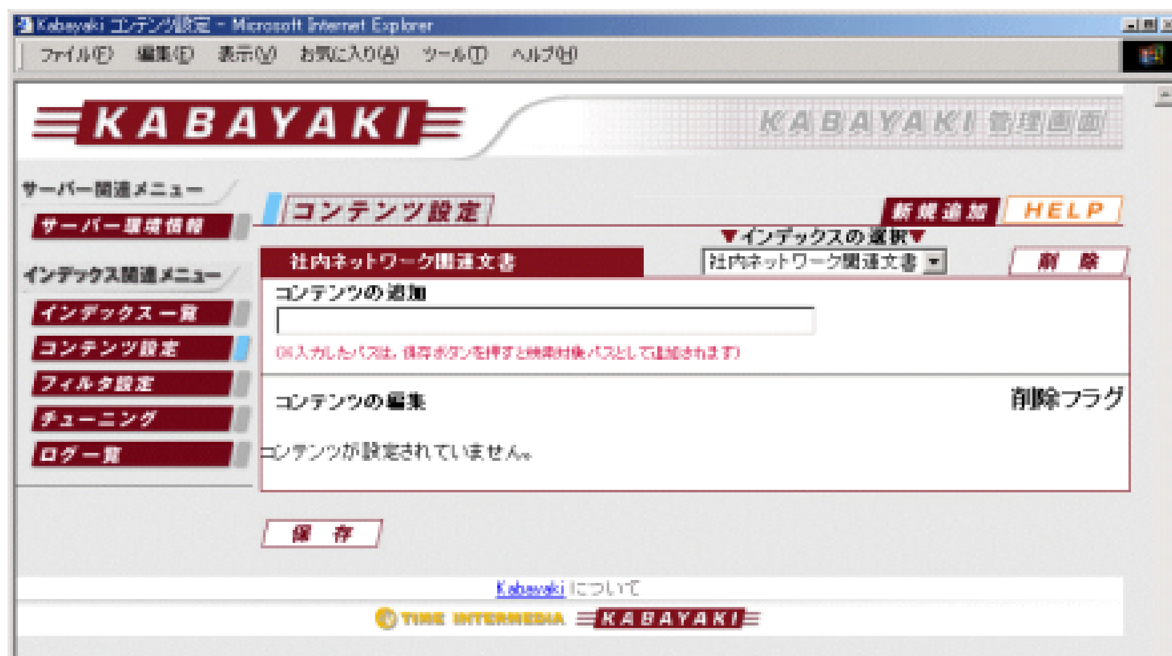
［登録］ボタン

入力されている内容で，Kabayaki へインデックスを新規に登録します．

コンテンツ設定

インデックスに対して，検索対象のドキュメントがある場所:コンテンツを，新たに追加・設定します．なお，コンテンツについては，『Kabayaki の管理とは コンテンツとインデックス』を参照してください．

メインメニューのコンテンツ設定をクリックするか，インデックス一覧の表示名もしくは内部名



をクリックすると，次のようなコンテンツ設定画面が表示されます．

画面に表示されている情報は以下のとおりです．

[コンテンツの追加]

入力フィールドに，検索の対象にしたい:インデクシングしたい文書が含まれるディレクトリのパスを入力します．パスは絶対パスで記述します．検索の対象となるファイルは，指定されたディレクトリ以下の全てのファイルとなります．

[コンテンツの編集]

コンテンツの一覧が表示されます．

Kabayaki をインストールした直後やコンテンツが一つも設定されていない場合は，

　コンテンツが設定されていません。

と表示されます．

コンテンツの追加フィールドには，検索の結果に表示されるドキュメントの場所を指定します．Kabayaki は，コンテンツ:検索対象のディレクトリパスの情報をそのまま保持するので，コンテンツを追加しただけでは，検索を行うユーザのウェブブラウザでアクセス可能な URL にならない場合があります．

例えば，/usr/local/doc/network に収められている文書を検索する場合には，まず，コンテンツの追加にディレクトリパスを入力して保存した後，コンテンツの編集で次のように変更して保存します．

例:

コンテンツの追加，へ入力して保存		コンテンツの編集，で変更して保存
http://usr/local/doc/network	->	**http://www.example.com/doc/network**

さて，WWW ブラウザで検索結果を正しく表示させるためには，さらに，文書が収められている WWW Server で設定しなくてはならないことがあります．
Apache HTTP Server を例に，その設定の方法を述べます．

Apache HTTP Server では，標準の設定ファイルである httpd.conf を編集することでおこないます．先の例なら，以下のように httpd.conf へ追記します．

```
Alias          /doc/network                    "/usr/local/doc/network"
```

コンテンツの追加で，異なるディレクトリパスを設定するたびに，上記のような Alias ディレクティブを追加してゆきます．
Alias を追記したら，Apache HTTP Server を再起動して，httpd.conf の変更を反映させます．
Apache HTTP Server の最新版では，以下のように Apache HTTP Server をインストールしたディレクトリへ移動して，コマンドを実行します．

```
# cd /usr/local/httpd/bin
# ./apachectl graceful
```

[削除フラグ] チェックボックス

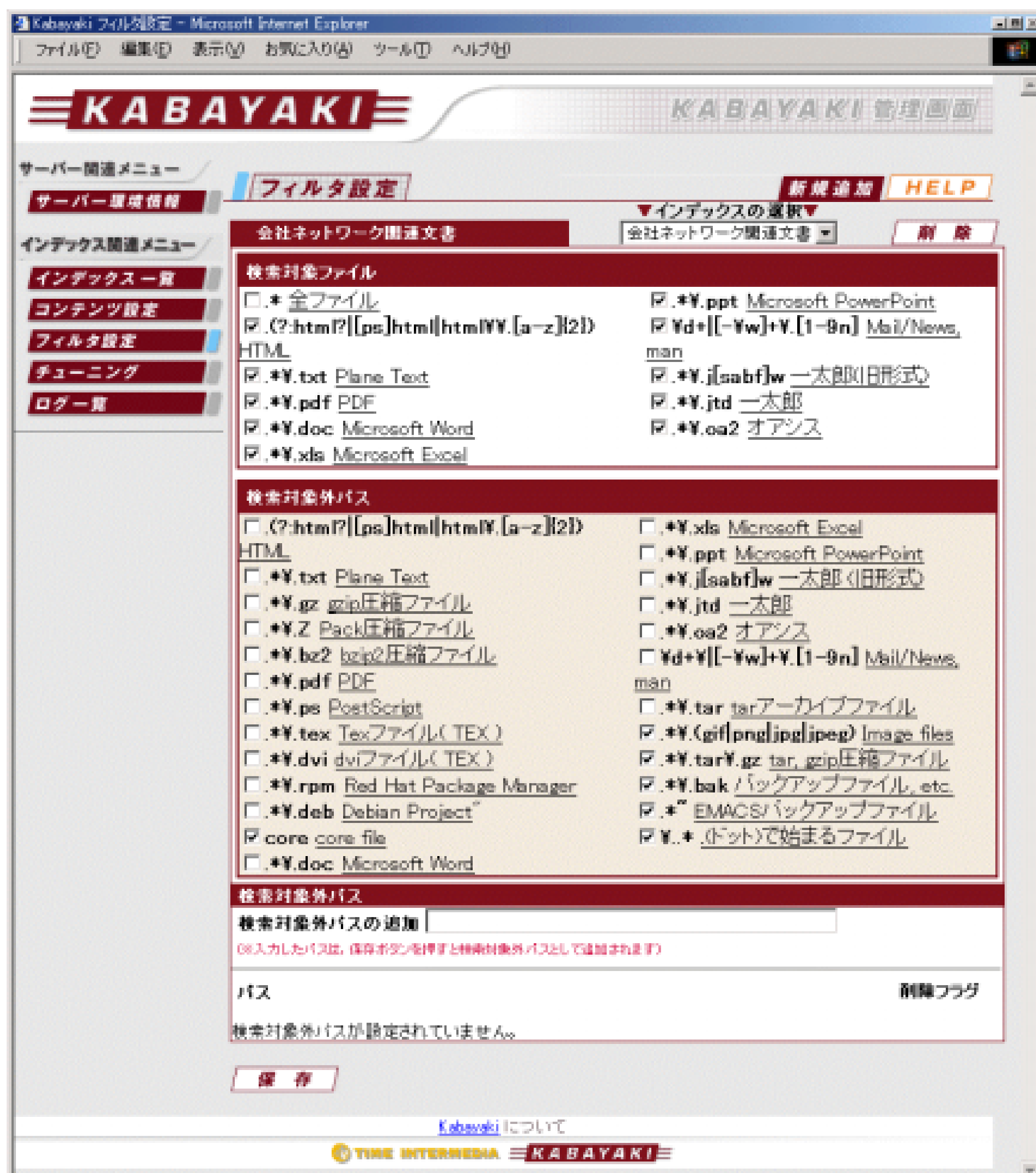
コンテンツを削除するには，コンテンツの編集の一覧の右側のチェックボックスをチェックし，[保存] ボタンをクリックすると，コンテンツが削除され，Kabayak の設定が変更されます．

フィルタ設定

検索の対象にするコンテンツを，ファイルの拡張子によるファイル形式やサブディレクトリに依って，選別:フィルタリングすることができます．フィルタリング設定を上手に利用することで，無駄なファイルのインデクシングを回避でき，インデクシング時間やホストの資源を節約できます．なお，設定項目の優先順位は次の通りです．

優先順位 低　　　　　　　　　　　　　　優先順位 高

検索対象ファイル ＜ 検索対象外ファイル ＜ 検索対象外パス



メインメニューのフィルタ設定をクリックすると，次のような画面が表示されます．

検索対象ファイル

検索の対象にしたいファイル名の拡張子を選び，チェックボックスをチェックします．全ファイルをチェックすると，コンテンツにある全てのファイルが検索の対象となります．

検索対象外ファイル

検索の対象にしたくないファイル名の拡張子を選び，チェックします．検索対象ファイルの設定で全てのファイルを検索対象としたとき，検索の対象にしたくないファイルを選択するときに役立ちます．

検索対象外パス

検索の対象にしないファイルまたはディレクトリのパスを指定します．コンテンツのパス以下にある，検索の対象に含めたくないサブディレクトリを指定します．

［保存］ボタン

［保存］ボタンをクリックすると，入力されているフィルタ設定を保存します．

チューニング

Kabayaki では，コンテンツとインデックスの設定の基づいて検索をおこないますので，日々変化するコンテンツのインデクシングを効率よく管理する必要があります．チューニングでは，インデクシングを行なう間隔やインデクシング時に用いるメモリ設定等を変更・表示します．チューニングを上手に利用することにより，ホストの性能に合わせて Kabayaki の運用をすることができます．



メインメニューのチューニングをクリックすると，次のような画面が表示されます．

インデクシング間隔

インデクシングを行なう間隔を設定・表示します．ここでは，実時刻は設定されません．実際にインデクシングが行なわれる時刻については，ホストのシステム管理者へお尋ねください．

インデクシング実行時刻

インデクシングを行なう時刻を設定・表示します．ここで，24 時間表記で，実際にインデクシングを行なわせたい時刻を設定します．インデックス毎に個別に設定できます．

[手動実行] ボタン

このボタンを押すと，ただちにインデクシングが開始されます．実際にインデクシングが行なわれる時刻が判らず，コンテンツをすぐに検索の対象としたい場合に用いるとよいでしょう．

使用メモリサイズの上限値

この値を大きくすると，インデクシングにかかる時間が短縮されることがあります．しかし，ホストの物理メモリが少ない場合等は，かえって遅くなることもありますから注意してください．搭載されている物理メモリの容量に関しては，ホストのシステム管理者にお尋ねください．

ファイルサイズの上限値

インデクシングの対象となるファイルの大きさです．この値よりも大きなファイルはインデクシングの対象にはなりせん．コンテンツに含まれる全てのファイルの大きさをよく精査して，検索からもれないように注意して設定してください．

テキストの上限値

インデクシングの対象となるファイルのテキスト部分の大きさです．この値よりもテキスト部分が大きなファイルはインデクシングされません．

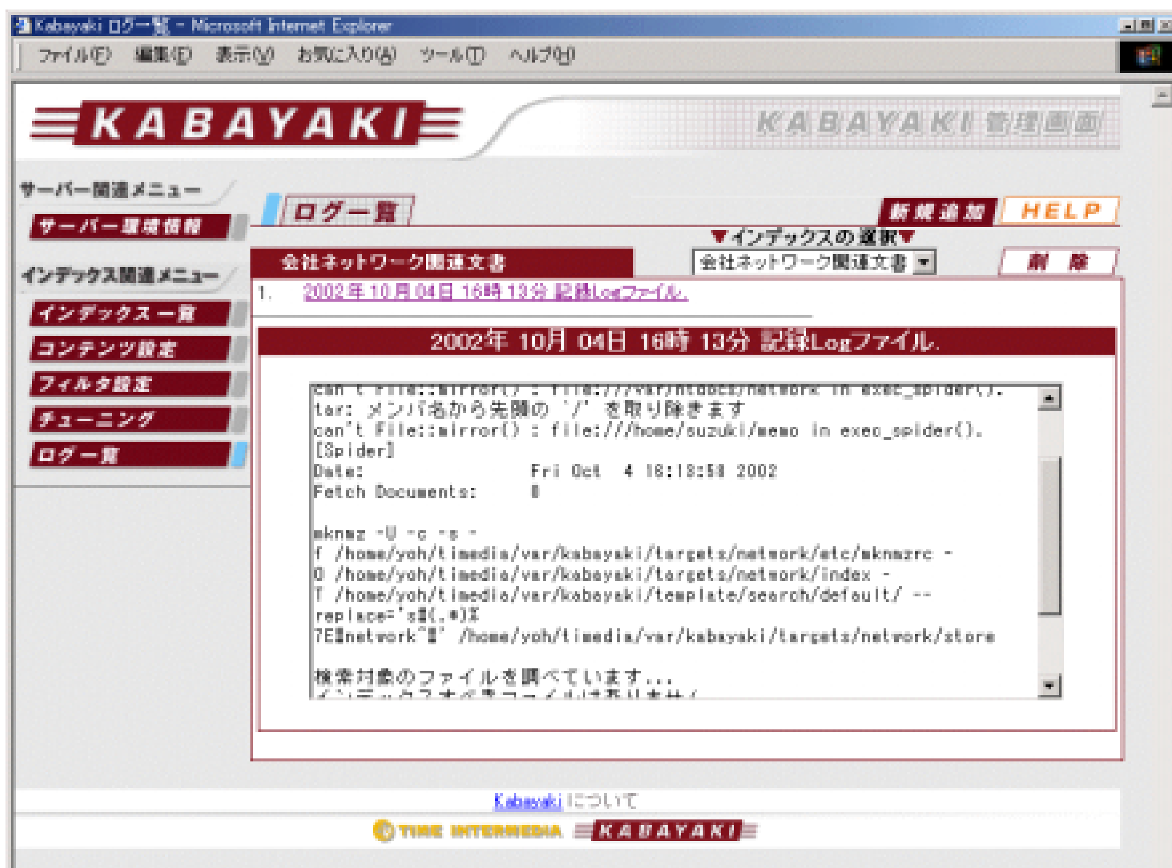
単語長の上限値

インデクシングの対象となる単語の長さです．この値よりも長い単語は，インデクシングされません．この場合，インデクシングされないのは単語のみで，その単語を含むファイルはインデクシングされます．

ログ一覧

インデクシングに関するログを表示します．ホストのシステム管理者は，ログを定期的に精査することによって，インデクシングのようすを確認することができます．

メインメニューのログ一覧をクリックすると，次のような画面が表示されます．



インデクシングのログの一覧が，番号 日時記録 Log ファイルという形式で表示されます．日付の若い順に並び替えらます．ログの一覧をクリックすると，下部のテキストエリアに記録されたログの詳細が表示されます．

詳細表示される項目には次の情報があります.

- 日付
- 追加された文書数
- 削除された文書数
- サイズ
- 更新された文書数
- 合計の文書数
- 追加キーワード数
- 合計キーワード数
- わかち書き
- 経過時間
- ファイル/秒
- システム
- Namazu
- エラー，警告，詳細

付録A
フィルタとプロパティ検索
文書フィルタとプロパティ検索詳細

文書フィルタ

Kabayaki は，文書フィルタと呼ばれる外部プログラムと連携して，HTML やテキストファイル以外の文書形式も検索の対象とすることができます．文書フィルタには次のような特長があります．

GPL 版

Namazu に添付されている文書フィルタを利用します．Xlhtml，wv 等の外部プログラムをインストールすることにより，Microsoft Word，Microsoft Excel などの文書ファイルを検索の対象とすることができます．詳しい情報については，Namazu に付属のドキュメントをご覧ください．

オプションパック sansyo をご購入いただきインストールしていただいた場合

Microsoft Word，Microsoft Excel，Microsoft PowerPoint，Adobe PDF について，GPL 版と比べ，より高い精度の検索を提供します．さらに，ジャストシステム 一太郎,富士通 OASYS 等，GPL 版では検索できない文書形式に対応します．

対応文書

GPL 版

namazu-2.0.10 と同等です．

オプションパック sansyo をインストールした場合

Text

HTML

Microsoft Word 95 / 97 / 98 / 2000 / 2002(XP)

Microsoft Excel 95 / 97 / 2000 / 2002(XP)

Microsoft PowerPoint 95 / 97 /2000 / 2002(XP)

ジャストシステム 一太郎 V5-V11

富士通 OASYS V6/V7,

Adobe PDF Acrobat 4.0 / 5.0

その他の文書については，namazu-2.0.10 と同等です．

プロパティ検索

プロパティとは，文書の属性情報で，文書を識別するためにファイルに埋め込まれた情報のことです．オプションパック sansyo をインストールしていただくと，プロパティを検索の対象とすることができます．プロパティには，タイトル，作成者の名前，キーワード，コメントなどを記録されています．プロパティに関しては，各文書を作成するアプリケーションの取扱説明書をご覧ください．

Kabayaki で検索できるプロパティは以下のとおりです．

プロパティ	説明	Word	Excel	PowerP	PDF	一太郎	OASYS
title	タイトル	○	○	○	○	×	×
author	著者	○	○	○	○	×	×
company	会社名	○	×	○	×	×	×

OLE オブジェクト検索

オプションパック sansyo では，OLE で埋め込まれたオブジェクトも検索の対象とすることができます．OLE に関しては，OLE をサポートしているアプリケーションの取扱説明書をご覧ください．

Kabayaki で検索できる OLE オブジェクトは以下のとおりです．

Dest／Source	Word	Excel	PowerPoint
Word	○	○	×
Excel	○	○	×
PowerPoint	×	×	×

オプションパック sansyo を Namazu の文書フィルタとしてお使いいただく場合について

オプションパック sansyo を Namazu の文書フィルタとして利用していただく場合は,namazu にてインデックスを作成する際に，Kabayaki の文書フィルタディレクトリ (/usr/share/kabayaki/filter)を指定してください．文書フィルタディレクトリの指定は，mknmzrc の$FILTERDIR で行います．

デフォルトの mknmzrc は/etc/namazu/mknmzrc に存在しますが,

mknmz -f /tmp/mknmzrc

等とコマンドを発行して，他の mknmzrc を読み込むこともできます．詳しくは Namazu のドキュメントをご覧ください